International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

Evaluate And Combine Several Audio Processing And Deep Learning Features For Heartbeat Sound Classification.

[1] Dr. R. Rambabu, [2]Korlana Sai Yaswanth Raj,[3]Challa Karthik, [4]Battula Chandu,

[1] Professor,Dept.of CSE, Rajamahendri Institute of Engineering & Technology, Bhoopalapatnam, Near Pidimgoyyi,Rajahmundry,E.G.Dist.A.P. 533107.

[2,3,4] Students, Dept.of CSE, Rajamahendri Institute of Engineering & Technology, Bhoopalapatnam, Near Pidimgoyyi,Rajahmundry,E.G.Dist.A.P. 533107.

ABSTRACT

The use of machine learning in healthcare has been on the rise. It is of utmost importance to address issues linked to heart-related statistics in light of the concerning number of fatalities worldwide caused by cardiovascular illnesses. How feature engineering affects classification performance is explored in this work. The following feature extraction methods were employed by a support vector machine: first, features extracted from audio signal processing; second, features extracted from a VGG-like architecture that had been pre-trained on Google's AudioSet; and lastly, features extracted from the ImageNet dataset that had been concatenated with features extracted from the VGG16 and VGG19 architectures. Last but not least, we used feature concatenation or majority vote to merge all methods. We compared our approaches to those in the literature and ran tests on two datasets from the PASCAL Classifying Heart Sounds Challenge. No matter the pre-training dataset, the experimental findings demonstrate that spectrograms used in deep learning and audio processing may potentially store the same important information for this application, and that more study is still advised.

Key Concepts—PASCAL, cardiac sound classification, feature engineering, music processing, DL, TL

1. INTRODUCTION

One of the main reasons people die all across the globe is heart disease. The World Health Organization and the American College of Cardiology report that cardiovascular illnesses account for over one-third of all deaths globally, which amounts to around 17.7 million people [1]. We must prioritize the early detection and treatment of cardiac problems if we want to reduce these numbers. The most economical method of listening to heart sounds—auscultation using a stethoscope—depends significantly on the doctor's ear sensitivity, expertise, and meticulous analysis for an accurate diagnosis. On the other hand, expert cardiologists have an accuracy rate four times higher than that of physicians in training, who can only manage an average of 20% [2]. Not only has it worsened with time, but it also raises expenses because to improper echocardiography orders, which is bad for patients since they can't get the treatment they need [4]. Because of this, the use of machine learning to problems involving the heart is gaining popularity. Another potential option for widespread and consistent first-level screening

of cardiac diseases may lie in society's digital use patterns, particularly with the rise of wearables. The PASCAL Network of Excellence sponsored the Classifying Heart Sounds Challenge, an audio data competition, in 2011 and 2012 [5]. Two datasets representing real-world scenarios, each with its own unique kind of background noise, made up the task. The assignment was split into two separate parts: segmenting heart sounds and classifying them. Only classification is addressed in this study. There are five distinct types. To have a normal class audio indicates a healthy heartbeat. At a heart rate below 140 beats per minute, a typical heart sound has a longer time between the "dub" and the "lub" sound, forming a distinct "lub dub, lub dub" pattern. Between S1 and S2 or S2 and S1 (but not on S1 or S2), the murmur class produces an acoustic signature like a "whooshing, roaring, rumbling, or turbulent fluid" sound. There are a lot of cardiac conditions that they could represent. An extra sound, such as a "lub-lub dub" or a "lub dub-dub," is used to identify audios from the extra heart sound class. Because ultrasonography has a hard time picking it up, finding it is crucial, even if it might be a sign of a problem or not. A wide range of sounds, from music to ambient noise, are presented in the audios of the artifact class. This is the most difficult to see, but finding it is crucial so the individual may take the test again. Records belonging to the extrasystole category include a heart sound that is not in sync with the rest of the recording, or what is effectively an extra heart sound that occurs from time to time but is not present consistently. Scientists have been trying to find ways to make the competition better for a long time now. Methods vary from optimizing model hyperparameters to using convolutional neural networks (CNN) to audio spectrograms, among others. In this research, we compare the efficacy of three distinct feature extraction methods: traditional features derived from audio processing, a CNN trained on audio data, and two convolutional neural networks (CNNs) trained on picture data. The plan is to evaluate them one by one and then merge them using feature concatenation or majority vote ensemble. Next, we evaluate the outcomes in relation to those earlier approaches, all the way up to the most recent publication that came to light during the time our trials were conducted.

2. METHODOLOGY

2.1. Datasets

The iStethoscope Pro iPhone app collected audio recordings from the general population and is part of Dataset A. By using features like real-time filtration and amplification, the software produces sound quality that is on par with, if not better than, commercially available digital stethoscopes, say cardiologists. In the Maternal and Fetal Cardiology Unit of the Real Hospital Portugũes (RHP) in Recife, Brazil, auscultations were recorded using the DigiScope Collector, which are included in Dataset B. Both Table 1 and Table 2 provide an overview of the dataset structure, including the number of files in each class label, as well as the sampling frequency and origin of those files.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

Table 1. Dataset A structure

Class	Quantity	Audio Information
Normal	31	iStethoscope (iPhone) 44.1 kHz
Murmur	34	
Extra Heart Sound	19	
Artifact	40	
Unlabeled	52	
Total	176	

Table 2. Dataset B structure

Class	Quantity	Audio Information
Normal	320	Digital Stethoscope 4 kHz
Murmur	95	
Extrasystole	46	
Unlabeled	195	
Total	656	

2.2. Audio Processing Features

Melfrequency cepstral coefficients (MFCCs), zero-crossings, spectral centroid, roll-off frequency, and chromagram were the metrics employed in audio signal processing. A chromagram is a projection of the audio spectrum onto the 12 semitones of the musical octave. Twenty MFCCs were among the twenty-four features we obtained after adding up the zero-crossings and averaging the spectral centroid, roll-off frequency, and chromagram values. Section 2.3: Deep Entries A 256-mel band spectrogram was produced using an FFT window of 2,048 samples, 512 samples between each frame, and an energy-based mel scale. In the end, values were transformed to the decibel (dB) scale to ensure no data was lost. We used spectrograms as inputs and retrieved deep learning features from the VGG16 and VGG19 models, which were pre-trained on the ImageNet dataset, from their second to last dense layer (fc1 or fc6) [6]. The input resolution of 224 × 224 × 3 was used to resize the spectrograms. We also used a CNN dubbed VGGish to extract deep learning features since its design is comparable to VGG's [7]. That one, on the other hand, comes pre-trained on Google's AudioSet, which is a database of 2,084,320 10-second audio snippets extracted from videos on YouTube and organized into 632 classes by humans [8]. Sec. 2.4. Classifiers
For the multi-class classification, we opted for the support vector machine (SVM) method because it is reliable, requires minimal training data, and has a track record of success with heartbeat sound categorization [9]. Every time we used the SVM in our method, we heuristically set its hyperparameters. The regularization parameter C could have values between 10−4 and 104, and the kernels could be either linear or radial basis function (RBF). The coefficient gamma could be either equal to the inverse of the number of features or the inverse of the number of features multiplied by its variance. The values that were ultimately utilized for each dataset are shown in Tables 3 and 4. The selection of models and tweaking of hyperparameters were not priorities.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

Table 3. SVM hyperparameters for Dataset A

Method	C	Kernel	Gamma
Audio Features	100	rbf	auto
VGGish	1	rbf	scale
VGG16+VGG19	0.001	linear	-
Feature Concatenation	1	linear	-

Table 4. SVM hyperparameters for Dataset B

Method	C	Kernel	Gamma
Audio Features	9	rbf	auto
VGGish	1	rbf	scale
VGG16+VGG19	5	rbf	auto
Feature Concatenation	0.001	linear	-

2.5. Evaluation Criteria

In order to compare our techniques to the other described alternatives, we used the metrics stated by the challenge to assess their efficacy. Accuracy, sensitivity, and specificity form its fundamental basis. In order to assess the diagnostic capabilities (i.e., the capacity to prevent failure) of various test methods, we compute the Youden's Index γ for both datasets.

$$\gamma = sensitivity - (1 - specificity) \qquad (1)$$

As an example, we take Dataset A's artifact class and apply Youden's Index to it. Then, we take Dataset B and apply it to the troublesome heartbeats class (murmur and extrasystole combined). However, we just evaluate the heart issue classes (murmur and additional heart sound combined) for calculating the F-Score for Dataset A, with β set to 0.9. Also, we just use Dataset B to calculate the discriminant power (DP), a measure of an algorithm's ability to distinguish between positive and negative examples:

$$DP = \frac{\sqrt{3}}{\pi}\left(\log\left(\frac{sensitivity}{1 - sensitivity}\right) + \log\left(\frac{specificity}{1 - specificity}\right)\right) \qquad (2)$$

An ineffective discriminant is indicated by a DP below 1. The algorithm is restricted if the DP is less than 2. If the DP is less than 3, it means the performance is fair. And it may be said to be an excellent algorithm in any other scenario. For samples including cardiac issues (including both murmur and extrasystole categories combined), the DP is computed. We used the assessment script, which was given by the challenge organizers as an Excel file, which included all of these metrics computations. Experiments (2.6) We ran three separate categorization algorithms separately. Using a support vector machine (SVM) classifier to extract ASP characteristics from audio recordings was the next step. The second one relied on spectrogram generation, feature extraction from VGGish (deep learning via transfer learning), and support vector machine (SVM) classification. The third one included creating spectrograms, sending them into the VGG16 and VGG19 simultaneously so that they may transfer learn—deep learning features from

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

their second-to-last layer, dense layer fc1 or fc6—and then using an SVM classifier to combine the features. In the end, we used an SVM classifier that had the characteristics of all three approaches integrated or a majority vote of the methods' predictions to combine them. Figure 1 shows the procedure that was explained. The open-source programming language Python was used for the implementations and experiments, with librosa, TensorFlow, and scikit-learn being the major tools used. The VGGish was used for transfer learning via their public GitHub repository. A total of 8,192 features (4,096 each) were generated by VGG16 and VGG19, compared to 128 features by VGGish. The total number of features was 8,344 when all three approaches' features were combined. After that, we used principal component analysis (PCA) to lower the dimensionality. With a total explained variance of 99.77% and 100 components for Dataset A and 400 components for Dataset B, respectively, we can see that the characteristics of both datasets were effectively reduced.

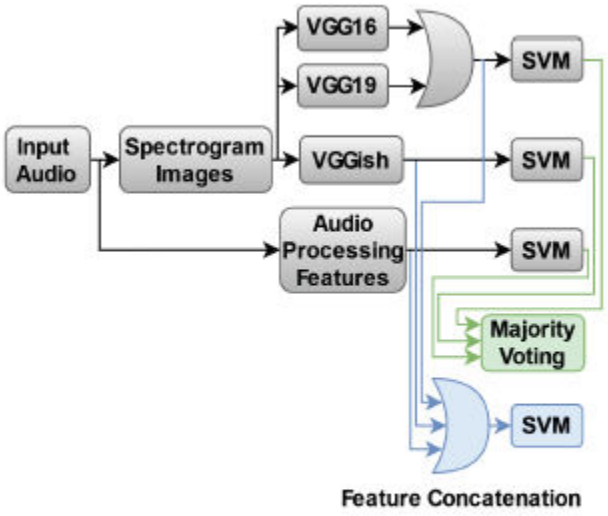

Fig. 1. Schematic diagram of our experiments.

3. RESULTS AND DISCUSSION

Table 5 displays the results obtained from all methods and the two combinations of them on Database A. Table 6 also includes these findings with those from other methodologies and publications, such as the official entries to the competition [13–20]. The results in [20] support our choice to merge VGG16 and VGG19, especially because of the notable improvement in extrasystole accuracy. Both Table 7 and Table 8 display the findings for Dataset B in a similar fashion. When it comes to modeling, there is no clear winner when it comes to feature extraction methods. Even their disparities did not hold true for the two sets of data. If we look at Dataset A, we can see that the classical audio features technique has the best cost-benefit ratio, but we can't say the same for Dataset B. There seems to be a fundamental difference between these two issues from a supervised learning perspective, even if Datasets A and B are identical in form and serve

very comparable purposes. Testing on a larger number of datasets, still from diverse sources but with the same goal classes, might provide a more accurate empirical assessment of this. Taken together, the results show that the best feature extraction techniques for one dataset may not be the best fit for another, thus it's best to experiment to find out what works. There is no guarantee that combining methods will lead to better results. Looking at the various criteria separately, VGG16+VGG19 got the most best scores on Dataset A, while VGGish had the highest overall precision score. When looking at overall accuracy, feature concatenation performed far worse on Dataset B than majority voting, which performed just slightly better. Spectrograms seem to retain all of the pertinent information from the original audio signal, and characteristics across these various approaches are more overlapping than complementing. However, storing and processing spectrograms is more costly. This use case may not need the high dimensionality often associated with visual activities. Almost all of the data was retained with a main component count ranging from 1% to 5% of all features. In spite of PCA's lack of intended use, it was able to improve the signal-to-noise ratio by decreasing the amount of background noise. From the standpoint of computing resources, this might be helpful for situations like efficient (re)training and feature storage, especially when working with constrained hardware. Dataset A is much smaller than Dataset B, however deep neural networks VGGish and VGG16+VGG19 outperformed audio processing characteristics. This highlights the efficacy of transfer learning, as it was not necessary for the dataset used for the downstream job to be as large as the one used for pre-training the models, or even large at all. The strategy could be essential depending on the objective. The achievement of a flawless score on the extrasystole precision, which has been traditionally challenging to categorize, was one of the most unexpected outcomes. This highlights the need of knowing what you want out of an optimization effort and keeping in mind which statistic is most relevant to your specific use case and needs. As an alternative to seeing it as a multiclass issue, it may suggest the use of a combination of models, with each model focusing on a different objective. Contrary to what was said, the pre-training dataset is not required to be in close proximity to the dataset of the downstream job. Compared to VGG16+VGG19, which was pretrained on image data, we anticipated that VGGish, which was trained on audio data, would perform better. This might indicate that after the spectrograms transform it into a visual task, the model's pattern-spotting abilities are the most important factor, as it was not the case for Dataset B and not universally for Dataset A. On Dataset A, neither the top scorer VGGish nor any combination of voters were able to outperform CNN-SVM, whereas majority voting did so on Dataset B. In order to do a more thorough comparison, it would be interesting to include more metrics into this study, such as the training and prediction runtime and memory consumption of each technique. If state-of-the-art performance is the goal, we anticipate transformer-based designs to achieve it, even if we also suggest extending this effort to concentrate on model optimization.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

4. CONCLUSIONS

Using feature engineering as a framework, we examined the categorization of heartbeat sounds in this article. We conducted our experiments on two difficult PASCAL Classifying Heart Sounds Challenge datasets. We evaluated all three of our strategies separately and in tandem using the same standards as the contest. Additionally, we contrasted them with prior efforts. Based on our findings, the feature space for this application could be much less than what is often seen in vision tasks. It is nevertheless important to incorporate classical audio processing qualities in the trade-off analysis, since they may perform better than first thought. It is not guaranteed that combining methods will result in optimal performance. It is nevertheless recommended to experiment with various ways while keeping clear and acceptable assessment criteria in mind. It seems that there is no need for the pre-training dataset to be identical to the downstream task, yet transfer learning still showed usefulness. And lastly, spectrograms seem to include all the pertinent data for this specific task, which opens up a world of possibilities since visual research is much more advanced than audio research, particularly when it comes to the availability of computer resources.

Our goal in doing this research was to provide useful information for creating and implementing early diagnostic tools for cardiac problems. What has been dubbed "the great consolidation" in machine learning is further supported by this, in our opinion.

Table 5. Results for Dataset A

Dataset A	Method				
Evaluation Criterion	**Audio Features**	**VGGish**	**VGG16+VGG19**	**Majority Voting**	**Feature Concatenation**
Precision of Normal	0.64	0.53	**0.73**	0.57	0.69
Precision of Murmur	**0.79**	0.71	0.77	0.77	0.77
Precision of Extrasound	0.71	**1.00**	0.50	0.80	0.50
Precision of Artifact	0.80	0.89	**1.00**	0.80	**1.00**
Sensitivity of Artifact	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
Specificity of Artifact	0.64	0.61	**0.67**	0.61	**0.67**
Sensitivity of Heart Problem	**0.73**	0.59	**0.73**	0.64	0.68
Precision of Heart Problem	0.76	0.76	0.64	**0.78**	0.65
Youden Index of Artifact	0.64	0.61	**0.67**	0.61	**0.67**
F-Score of Heart Problem	**0.33**	0.30	0.30	0.32	0.30
Total Precision	2.94	**3.13**	3.00	2.94	2.96

Values in bold are the best scores among methods for each criterion.

Table 6. Comparison of obtained results with other methods on Dataset A

Dataset A	Previous Methods in the Literature									Our Combined Methods	
Evaluation Criterion	**J48 [13,14]**	**MLP [13,14]**	**CS-UCL [13,15]**	**SS [16]**	**SS-PLSR [16]**	**2D-PCA [17]**	**SS-TD [18]**	**SVM-DM [19]**	**CNN-SVM [20]**	**Majority Voting**	**Feature Concatenation**
Precision of Normal	0.25	0.35	0.46	0.67	0.60	0.56	0.67	0.62	0.59	0.57	**0.69**
Precision of Murmur	0.47	0.67	0.31	0.91	0.91	0.91	**1.00**	**1.00**	0.77	0.77	0.77
Precision of Extra Heart Sound	0.27	0.18	0.11	0.37	0.44	0.30	0.43	**1.00**	0.83	0.80	0.50
Precision of Artifact	0.71	0.92	0.58	0.76	0.94	0.94	0.80	0.64	**1.00**	0.80	**1.00**
Sensitivity of Artifact	0.63	0.69	0.44	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
Specificity of Artifact	0.39	0.44	0.44	0.58	0.64	0.58	0.64	0.58	**0.69**	0.61	0.67
Youden Index of Artifact	0.01	0.13	-0.09	0.58	0.64	0.58	0.64	0.58	**0.69**	0.61	0.67
F-Score of Heart Problem	0.20	0.20	0.14	0.28	0.30	0.26	0.30	0.31	**0.33**	0.32	0.30
Total Precision	1.71	2.12	1.47	2.71	2.89	2.80	2.90	3.17	**3.19**	2.94	2.96

Values in bold are the best scores among methods for each criterion.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

Table 7. Results for Dataset B

Dataset B	Method				
Evaluation Criterion	Audio Features	VGGish	VGG16+VGG19	Majority Voting	Feature Concatenation
Precision of Normal	0.76	0.76	**0.78**	0.77	**0.78**
Precision of Murmur	0.61	0.85	0.79	**0.86**	0.79
Precision of Extrastole	0.50	0.50	**1.00**	**1.00**	0.50
Sensitivity of Heart Problem	**0.34**	0.31	**0.34**	0.32	**0.34**
Specificity of Heart Problem	0.90	0.97	0.97	**0.98**	0.96
Youden Index of Heart Problem	0.24	0.28	**0.31**	0.30	0.30
Discriminant Power	0.38	0.64	0.68	**0.73**	0.62
Total Precision	1.87	2.11	2.57	**2.63**	2.07

Values in bold are the best scores among methods for each criterion.

Table 8. Comparison of obtained results with other methods on Dataset B

Dataset B	Previous Methods in the Literature									Our Combined Methods	
Evaluation Criterion	J48 [13, 14]	MLP [13, 14]	CS-UCL [13, 15]	SS [16]	SS-PLSR [16]	2D-PCA [17]	SS-TD [18]	SVM-DM [19]	CNN-SVM [20]	Majority Voting	Feature Concatenation
Precision of Normal	0.72	0.7	0.77	0.74	0.76	0.78	**0.83**	0.77	0.81	0.77	0.78
Precision of Murmur	0.32	0.3	0.37	0.66	0.65	0.57	0.7	0.76	0.76	**0.86**	0.79
Precision of Extrasystole	0.33	0.67	0.17	0.24	0.33	0.23	0.15	0.5	0.56	**1.00**	0.50
Sensitivity of Heart Problem	0.22	0.19	0.51	0.24	0.34	0.41	0.49	0.34	**0.54**	0.32	0.34
Specificity of Heart Problem	0.82	0.84	0.59	0.84	0.9	0.84	0.84	0.95	0.91	**0.98**	0.96
Youden Index of Heart Problem	0.04	0.02	0.01	0.13	0.24	0.24	0.33	0.29	**0.45**	0.30	0.30
Discriminant Power	0.05	0.04	0.09	0.24	0.36	0.3	0.39	0.54	**0.6**	0.73	0.62
Total Precision	1.37	1.67	1.31	1.57	1.75	1.58	1.68	2.03	2.15	**2.63**	2.07

Values in bold are the best scores among methods for each criterion.

5. REFERENCES

[1] G. A. Roth, C. Johnson, A. Abajobir, et al., "Global, regional,and national burden of cardiovascular diseases for 10 causes,1990 to 2015," Journal of the American College of Cardiology,vol. 70(1), pp. 1–25, 2017.

[2] S. Mangione and L. Z. Nieman, "Cardiac auscultatory skills ofinternal medicine and family practice trainees: a comparisonof diagnostic proficiency," JAMA, vol. 278(9), pp. 717–722,1997.

[3] E. Etchells, C. Bell, and K. Robb, "Does this patient have anabnormal systolic murmur?," JAMA, vol. 277, pp. 564–571,1997.

[4] U. Alam, O. Asghar, S. Q. Khan, S. Hayat, and R. A. Mali,"Cardiac auscultation: an essential clinical skill in decline,"The British Journal of Cardiology, vol. 17, pp. 8–10, 2010.

[5] P. Bentley, G. Nordehn, M. Coimbra, andS. Mannor, "The PASCAL Classifying HeartSounds Challenge 2011 (CHSC2011) Results,http://www.peterjbentley.com/heartchallenge/index.html.

[6] J. Deng, W. Dong, R. Socher, K. Li L.-J. Li, and L. Fei-Fei,"Imagenet: A large-scale hierarchical image database," in Conferenceon Computer Vision and Pattern Recognition. IEEE, 2009, p. 248–255.

[7] S. Hershey et al., "CNN architectures for large-scale audioclassification," in International Conference on Acoustics,Speech and Signal Processing (ICASSP). IEEE, 2017, pp. 131–135.

[8] J. F. Gemmeke et al., "Audio set: An ontology and humanlabeleddataset for audio events," in International Conferenceon Acoustics, Speech and Signal Processing (ICASSP). IEEE,2017, pp. 776–780.

[9] W. Zhang, J. Han, and S. Deng, "Heart sound classificationbased on scaled spectrogram and partial least squares regression,"Biomedical Signal Processing and Control, vol. 32, pp.20–28, 2017.

[10] B. McFee, C. Raffel, D. Liang, et al., "librosa: Audio andmusic signal analysis in python," in 14th python in scienceconference, 2015, pp. 18–25.

[11] M. Abadi, A. Agarwal, P. Barham, et al., "Tensorflow: Largescalemachine learning on heterogeneous distributed systems,"http://download.tensorflow.org/paper/whitepaper2015.pdf, 2015.

[12] Agrawal, K. K. ., P. . Sharma, G. . Kaur, S. . Keswani, R. . Rambabu, S. K. . Behra, K. . Tolani, and N. S. . Bhati. "Deep Learning-Enabled Image Segmentation for Precise Retinopathy Diagnosis". *International Journal of Intelligent Systems and Applications in Engineering*, vol. 12, no. 12s, Jan. 2024, pp. 567-74, https://ijisae.org/index.php/IJISAE/article/view/4541.

[13] E. F. Gomes, P. J. Bentley, M. Coimbra, E. Pereira, andY. Deng, "Classifying heart sounds: Approaches to the pascalchallenge," in International Conference on Health Informatics,2013, p. 337–340.

[14] E. F. Gomes and E. Pereira, "Classifying heart sounds usingpeak location for segmentation and feature construction," inWorkshop Classifying Heart Sounds, 2012, p. 480–492.

[15] Y. Deng and P. J. Bentley, "A robust heart sound segmentationand classification algorithm using wavelet decomposition andspectrogram," in Workshop Classifying Heart Sounds, 2012, p. 1–6.

[16] S. Deng and J. Han, "Towards heart sound classificationwithout segmentation via autocorrelation feature and diffusionmaps," Future Generation Computer Systems, vol. 60, pp.13–21, 2016.

[17] Samota, H. ., Sharma, S. ., Khan, H. ., Malathy, M. ., Singh, G. ., Surjeet, S. and Rambabu, R. . (2024) "A Novel Approach to Predicting Personality Behaviour from Social Media Data Using Deep Learning", *International Journal of Intelligent Systems and Applications in Engineering*, 12(15s), pp. 539–547. Available at: https://ijisae.org/index.php/IJISAE/article/view/4788

[18] L. D. Avendano-Valencia, J. I. Godino-Llorente, M. Blanco-Velasco, and G. Castellanos-Dominguez, "Feature extractionfrom parametric time-frequency representations for heart murmurdetection," Annals of Biomedical Engineering, vol. 38(8),pp. 2716–2732, 2010.

[19] S. C. Oliveira, E. F. Gomes, and A. M. Jorge, "Heart soundsclassification using motif based segmentation," in 18th InternationalDatabase Engineering & Applications Symposium.Association for Computing Machinery, 2014, p. 370–371.

[20] W. Zhang, J. Han, and S. Deng, "Heart sound classificationbased on scaled spectrogram and tensor decomposition," ExpertSystems with Applications, vol. 84, pp. 220–231, 2017.

[21] F. Demir, A. Sengur, V. Bajaj, et al., "Towards the classificationof heart sounds based on convolutional deep neural network,"Health Information Science and Systems, vol. 7(16), 2019.